
Final copy appears in *Journal of the History of the Behavioral Sciences* 50:3, Summer 2014, pp. 323-324

For many historians of the behavioral sciences, 'testing' is taken to mean aptitude testing. But we can also learn from histories of achievement testing. Although measuring something qualitatively different than aptitude – knowledge acquired; skills learned; industriousness - achievement testing also requires standardization. Historians of the behavioral sciences can learn how uniform conditions were created which made it possible to collect meaningful statistics. The standardized achievement examination is just as much a statistical survey as a pedagogical tool. These considerations mean that although William J. Reese's *Testing Wars in the Public Schools* is mainly a history of education, it is also a history of testing and measurement. Indeed Harvard University Press does advertise *Testing Wars* as belonging to this category, and so I will review it as such a history.

Reese closely studies Boston's schools and educational milieu before and after the introduction of written achievement examinations in 1845. The subtitle of the book is that it is *a Forgotten History*, highlighting the point that we overlook how written and "standardized" examinations have raised hackles since their inception. Indeed Reese argues that this initial controversy between educational activists and school teachers "set the basic parameters for debates about competitive tests and their meaning that echo" to the present day (Reese, *Testing Wars*, p. 5). By studying the history of this first dispute, we learn how some of the themes of this 19th century debate crystallized and came to be endlessly repeated in present-day America.

The format of Reese's book is chronological, a narrative focusing on a bitter dispute between reformers Horace Mann and Samuel Gridley Howe against Boston's grammar school masters. The two men sought to change the state of these schools by introducing a uniform and written achievement exam posing the same questions about history, geography, definitions and grammar. It is probably not accidental that both men were committed phrenologists, and perhaps a little more attention might have been paid to this striking affinity. But the great strength of *Testing Wars* is its extremely detailed discussion of the political maneuverings before and after the 1845 examination, information Reese has gleaned from the various pamphlets which blanketed Boston, either favoring or denouncing the test. The claims and counterclaims in these pamphlets are familiar to us now, but only because it was in them that such claims were first enunciated: either exams overly routinized education, focusing attention only upon the act of passing exams - or they generated useful knowledge about the actual state of Boston's grammar schools.

Reese shows how this 1845 exam had a great impact upon American education, but notes that this only came later. At first the big test did not seem to have much of an effect. Some awful teachers - given to flogging and drinking beer in front of their students - lost their jobs, but mainly because of parents' complaints (p. 123). This delayed impact seems to have been because exams provided permanent records of work, unlike the evanescent oral "exhibitions", where hand-picked students recited aloud their rehearsed answers to public questions. Not only could these inscriptions testify to individual performances – they could
be aggregated, and provide statistics. When arrayed in temporal sequence, examinations could be used to demonstrate some sort of pattern. Written tests did not only work, then – they were also seen to work, showing "reform" in action. That said, reform means losing something as well - I would have liked a little more information about what was lost when achievement testing shifted from spoken exhibitions to written examinations. Andrew Warwick's *Masters of Theory*, for instance, shows how mathematics, not a field one would normally associate with oratory, changed when its mode of assessment shifted away from oral debate.

I noted that Reese's work pays very close attention to the primary sources, and is made livelier for it. But I do wish it had stepped back periodically to engage with secondary sources and been more epistemologically curious. Histories of measurement and quantification are duly cited, such as Theodore Porter's *Trust in Numbers*, but such footnoting seems a little ritualistic. For *Testing Wars* wrestles very little with the points raised by such works: it perfunctorily notes that the usual "subjective standards" were applied (p. 70) or that "Bostonians based their judgments about schools on impressions more than anything else" (p. 40). Setting aside fraught discussions about subjectivity and objectivity provided by "impressions," I wanted more answers as to why quantification and statistics become favored tools for the school reformer and why mere impressions became insufficient. Why did Mann consider it important to have a "daguerreotype likeness" – that is, a photograph – of the "pupils' minds" (p. 143)? What is it about numerical scores that can act as a decontextualized symbol for information, and then go on to be publicized by would-be reformers who see these scores as meaningful evidence? It is all too easy to point out what a book missed. But I bring these questions up because they would seem to be necessary to fulfill the author's own stated goal of studying how the parameters for debates about testing emerged (p. 5). For in order to understand such parameters, we need to probe the taken-for-granted ways in which examinations actually work.

James Elwick, Science and Technology Studies, York University
